2 Statistics, Data Analysis, and Probability

14 questions	12 questions	17 questions	17 questions	8 questions	12 questions
Number Sense	Statistics, Data Analysis, Probability	Algebra and Functions	Measurement and Geometry	Math Reasoning	Algebra I

CAHSEE contains 12 Statistics, Data Analysis, and Probability items. To demonstrate knowledge and skills in the Statistics, Data Analysis, and Probability strand, students must understand the fundamental concepts involved in data collection, display, and analysis. Students will be asked to determine ways to collect, organize, and display relevant data to answer questions; formulate questions that can be addressed with data; select and use appropriate statistical methods to analyze data; and develop and evaluate inferences and predictions that are based on data. In addition, students are required to understand and apply the basic concepts of probability.

Specifically, the standards in the Statistics, Data Analysis, and Probability strand include the following knowledge and skills:
- finding measures of central tendency to characterize data
- interpreting and evaluating conclusions based on data
- organizing and representing possible outcomes for events and expressing theoretical probabilities
- representing probabilities as ratios, proportions, and percents
- understanding the numerical continuum of probability between impossibility (0) and certainty (1)
- recognizing the difference between independent and dependent events
- displaying data appropriately, including both one- and two-variable data sets

The seven California academic content standards covered by the CAHSEE Statistics, Data Analysis, and Probability strand are discussed in the following pages.

| Strand | **Statistics, Data Analysis, and Probability (PS)** | **Rico's first three test scores in biology were 65, 90, and 73. What was his mean score?** |

| Standard | **6PS1.1** |
| **Compute the ~~range~~ mean, median, and mode of data sets.*** | |

3 test questions

Rico's first three test scores in biology were 65, 90, and 73. What was his mean score?

A 65

B 73

C 76

D 90

M02247

One of the major objectives of the Statistics, Data Analysis, and Probability strand is to give students tools to help them understand the uses and misuses of statistics. This CAHSEE content standard has three components: computation of the mean, computation of the median, and recognition of the mode of data sets. Statistical measures of central tendency represent important methods for summarizing and comparing single-variable data sets. Students should understand the significance of each as a measure of central tendency as well as the differences among these measures. For this standard, students will not be asked to find the median of an even number of values.

Students should know that:
- the median is the middle score of an ordered set of numbers, where half the scores are greater than the median and half are less.
- the mode is the number which appears most frequently.
- the mean is most affected by extreme values.

Sample Test Question
The correct answer is choice C. Students should recognize that they should compute the mean by first finding the sum $(65 + 90 + 73 = 228)$ and then dividing by $3 (228 \div 3 = 76)$.

Analysis of Distractors

Distractor A: minimum of the data set

Distractor B: median of the data set

C: correct answer

Distractor D: maximum of the data set

* The CAHSEE test blueprint does not include the crossed-out portion of this content standard.

Strand **Statistics, Data Analysis, and Probability (PS)**

Standard **6PS2.5**
Identify claims based on statistical data and, in simple cases, evaluate the validity of the claims.

1 test question

The Smithburg town library wanted to see what types of books were borrowed most often.
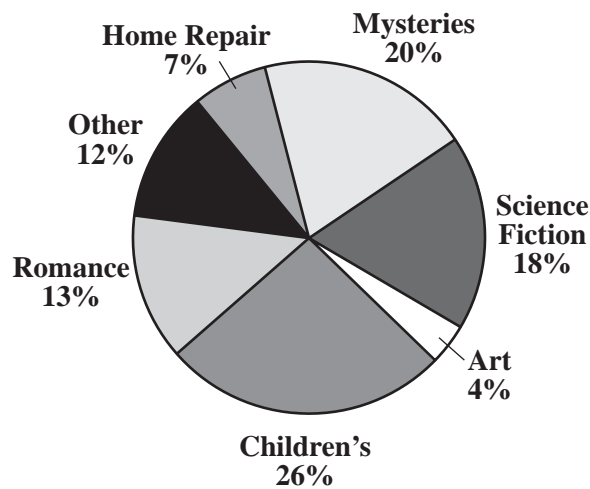
According to the circle graph shown above—

A more Children's books were borrowed than Romance and Science Fiction combined.

B more than half of the books borrowed were Children's, Mysteries, and Art combined.

C more Mysteries were borrowed than Art and Science Fiction combined.

D more than half of the books borrowed were Romance, Mysteries, and Science Fiction combined.

M02131

This content standard has two components: identifying claims made on the basis of statistical data and evaluating the validity of the claims based on statistical data. Because students should be able to understand statistical claims as well as they understand purely verbal arguments for or against a position, students should develop skills to evaluate the quality of data and conclusions based on data. CAHSEE questions for this standard may ask students to identify a valid claim based on data or to recognize a question for which the data could be used to provide an answer.

Sample Test Question
The correct answer is D. Students should be able to add the percentages from Romance, Mysteries, and Science Fiction and get 51%, thus making option D the correct one.

Analysis of Distractors

Distractor A: The percentage of Children's books (26) is not more than the sum of the percentages of Romance and Science Fiction $(13+18)$.

Distractor B: The sum of the percentages of Children's, Mysteries, and Art $(26+20+4)$ is exactly 50, which is not more than half.

Distractor C: The percentage of Mysteries (20) is not more than the sum of the percentages of Art and Science Fiction $(4+18)$.

D: correct answer

Strand **Statistics, Data Analysis, and Probability (PS)**

Standard **6PS3.1**
Represent all possible outcomes for compound events in an organized way (e.g., tables, grids, tree diagrams) and express the theoretical probability of each outcome.

1 test question

To get home from work, Curtis must get on one of the three highways that leaves the city. He then has a choice of four different roads that lead to his house. In the diagram below, each letter represents a highway, and each number represents a road.

Highway

Road	A	B	C
1	A 1	B 1	C 1
2	A 2	B 2	C 2
3	A 3	B 3	C 3
4	A 4	B 4	C 4

If Curtis randomly chooses a route to travel home, what is the probability that he will travel Highway B and Road 4?

A $\dfrac{1}{16}$

B $\dfrac{1}{12}$

C $\dfrac{1}{4}$

D $\dfrac{1}{3}$

M02512

Organizing structures, such as sample spaces, diagrams, and tables, are useful for the representations of probabilities, and the ability to create a structured representation of a complex situation is an important reasoning tool. To demonstrate achievement in this standard, students must recognize appropriate and correct representations of events. From the correct representation, they must derive an understanding of the relationship between the frequency of the outcome and its numerical expression. They should be able to determine a theoretical probability of any particular outcome based on a correct representation.

Sample Test Question

The correct answer is choice B. Students should recognize that if there are *n* possible outcomes for an independent event and *r* possible outcomes for another independent event, there are *nr* outcomes for the two events together. In other words, if there are *n* ways to do one thing and *r* ways to do another thing, there are *nr* ways to do the two things together. Thus, three highways times four roads equals 12 routes, and the probability of selecting any individual route is $\frac{1}{12}$.

Analysis of Distractors

Distractor A: squared the probability of taking one road

   B: correct answer

Distractor C: the probability of taking any one road

Distractor D: the probability of taking any one highway

Strand **Statistics, Data Analysis, and Probability (PS)**

Standard **6PS3.3**
Represent probabilities as ratios, proportions, decimals between 0 and 1, and percentages between 0 and 100 and verify that the probabilities computed are reasonable; know that if *P* is the probability of an event, 1-*P* is the probability of an event not occurring.

2 test questions
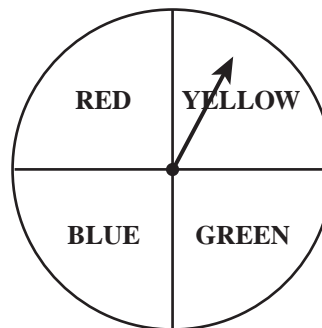
The spinner shown above is fair. What is the probability that the spinner will NOT stop on red if you spin it one time?

A $\dfrac{1}{4}$

B $\dfrac{1}{3}$

C $\dfrac{3}{4}$

D $\dfrac{4}{3}$

M00094

All students should understand that mathematical probability is used to predict what might happen in the future and that probabilities are ratios determined by considering the likely results or outcomes of events. CAHSEE questions for this standard cover all of the components of the standard. Students are expected to:

- know that probabilities are ratios that can be expressed as fractions, decimals, or percentages.
- compute the probability of a described event.
- verify the reasonableness of a computed probability.
- compute the probability that an event will not occur.

Sample Test Question
The correct answer is choice C. Students should recognize that because the four regions covered

by the spinner have equal areas, the probability of the spinner stopping on any one region is $\dfrac{1}{4}$.

Therefore, the probability that it will **not** stop on a given region is $1-\dfrac{1}{4}$ or $\dfrac{3}{4}$.

Analysis of Distractors

Distractor A: the probability that the event will occur (this value was not subtracted from 1)

Distractor B: misconception—only three quadrants were used to calculate the probability; the probability of the spinner landing on one of the three quadrants

C: correct answer

Distractor D: used a fraction to represent the total number of quadrants divided by three quadrants (all but red) or obtained the correct answer but then inverted it

Strand	**Statistics, Data Analysis, and Probability (PS)**

Standard   **6PS3.5**
Understand the difference between independent and dependent events.

1 test question

A bag contained four green balls, three red balls, and two purple balls. Jason removed one purple ball from the bag and did NOT put the ball back in the bag. He then randomly removed another ball from the bag. What is the probability that the second ball Jason removed was purple?

A $\dfrac{1}{36}$

B $\dfrac{1}{9}$

C $\dfrac{1}{8}$

D $\dfrac{2}{9}$

M03097

Discerning the difference between dependent and independent events is important in evaluating probabilistic outcomes. CAHSEE questions in this content standard require students to understand that events are independent of each other if the occurrence or non-occurrence of one event does not affect the probability of the occurrence or non-occurrence of another event. Similarly, students must recognize that events are dependent if the occurrence or non-occurrence of one event affects the probability of the occurrence or non-occurrence of another event. Computation may be required to determine the result of the independent or dependent events.

Sample Test Question
The correct answer is choice C. Students should determine that initially the bag contains 4 green, 3 red, and 2 purple balls, for a total of 9 balls. When 1 purple ball is removed and not replaced, the bag contains a total of 8 balls. Since there are now 8 balls with only 1 being purple, the probability of randomly choosing the purple ball is $\dfrac{1}{8}$.

Analysis of Distractors

Distractor A: the probability of randomly choosing both purple balls

Distractor B: the probability of randomly choosing 1 purple ball from 9 balls

       C: correct answer

Distractor D: added the two probabilities for an independent event $\left(\dfrac{1}{9}+\dfrac{1}{9}\right)$

Strand **Statistics, Data Analysis, and Probability (PS)**

Standard **7PS1.1**
Know various forms of display for data sets, ~~including a stem-and-leaf plot or box and-whisker plot;~~ use the forms to display a single set of data or to compare two sets of data.*

2 test questions

Speed of Four Runners in a 100-Meter Dash
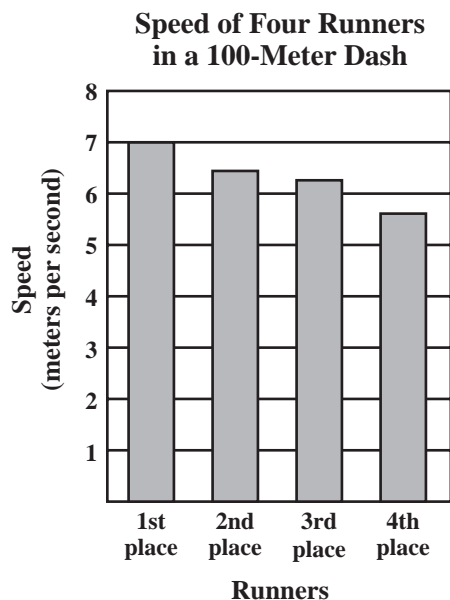
Based on the bar graph shown above, which of the following conclusions is true?

A   Everyone ran faster than 6 meters per second.

B   The best possible rate for the 100-meter dash is 5 meters per second.

C   The first-place runner was four times as fast as the fourth-place runner.

D   The second-place and third-place runners were closest in time to one another.

M00279

Large data sets are difficult to grasp mentally without an accessible visual representation. CAHSEE questions in this standard require students to recognize and interpret various forms of display and to compare two sets of data displayed the same way. The forms of display for single-variable data sets assessed on the CAHSEE include bar graphs, line graphs, scatterplots, pictographs, circle graphs, and Venn diagrams. CAHSEE questions for this standard may also require students to select an appropriate type of data display. Items for this standard may also focus on the students' ability to extract and manipulate data from a data display for use in a computation.

---

*   The CAHSEE test blueprint does not include the crossed-out portion of this content standard.

Sample Test Question
The correct answer is choice D. Students should use the graph to determine the speed of each runner and then evaluate and compare the four answer choices. From the graph, the first-place runner's speed was approximately 7 meters per second (m/s); the second-place runner's speed was approximately 6.5 m/s; the third-place runner's speed was approximately 6.3 m/s; and the fourth-place runner's speed was approximately 5.6 m/s. This data display shows that the second-place and third-place runners' times were closest together.

Analysis of Distractors

Distractor A: incorrect because the fourth-place runner ran at approximately 5.6 m/s

Distractor B: incorrect because all four runners ran faster than 5 m/s

Distractor C: incorrect because 7 m/s is not four times faster than 5.6 m/s

      D: correct answer

Probability Statistics

Probability Statistics

Strand	**Statistics, Data Analysis, and Probability (PS)**

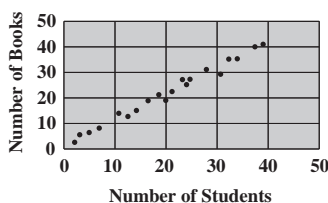**Which scatterplot shows a negative correlation?**

Standard **7PS1.2**
Represent two numerical variables on a scatterplot and informally describe how data points are distributed and any apparent relationship that exists between the two variables (e.g., between time spent on homework and grade level).
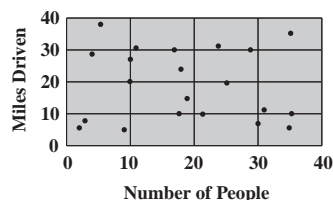
2 test questions

A

Number of Books vs Number of Students

C

Miles Driven vs Number of People

B

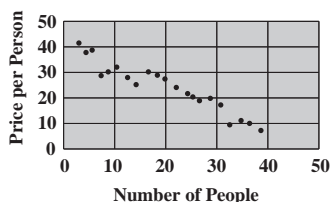Price per Person vs Number of People
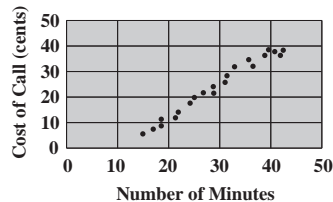
D

Cost of Call (cents) vs Number of Minutes

M02546

The identification of patterns and relationships, including clustering and trends, as well as the concept of correlation (positive, negative, or none) are significant aspects of using data. Students should understand correlation as a measure of the relationship between two variables, with negative correlation as the association of an increase in the value of one variable with a decrease in the corresponding value of the second variable. Students should also recognize that positive correlation is the association of an increase in the value of one variable with an increase in the corresponding value of the second variable.

CAHSEE questions for this standard address the following components of the standard:
- representing two variables on a scatterplot
- determining the distribution of the variables
- recognizing the general relationships between the two variables represented.

Sample Test Question
The correct answer is choice B: As the price per person increases, the number of people decreases, indicating a negative correlation.

Analysis of Distractors

Distractor A: represents a positive correlation

B: correct answer

Distractor C: represents no correlation

Distractor D: represents a positive correlation